DOCUMENT RESUME

ED 062 608 AC 012 576

AUTHOR Blaney, John P.
TITLE The Application of Formative and Summative Evaluation Techniques to Short Term Adult Education Courses.
PUB DATE 72
NOTE 27p.; Paper presented at the 1972 Adult Education Research Conference (Chicago, Ill.); Adapted in part from "A Guide to Developing and Evaluating Instruction in Continuing Education: Basic Concepts and Procedures"

EDRS PRICE MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS *Adult Education; Course Content; Discussion Experience; *Evaluation Techniques; *Instructional Materials; Literature; Objectives; Problem Solving; *Program Evaluation; *Projects; Response Mode

ABSTRACT

This paper identifies a problem concerning the practice of adult education, and describes one person's response to that problem. The problem is the lack of systematic evaluation of instruction in the field of adult education, and, more particularly, the absence of sufficient evaluation literature that is both useful and practical to those who earn their living in continuing education. The response is a developmental project aimed at producing for practitioners a set of materials that can be used in developing and evaluating instruction. The nature of these materials and how some practitioners have been trained to use them are described. The object of the paper is to generate discussion about whether the identified problem is real and accurately described, and how this particular response to it, if adequate, might be improved. (Author/CK)

The Application of Formative and Summative Evaluation Techniques to Short Term Adult Education Courses

a presentation to

1972 Adult Education Research Conference
Chicago, Illinois

John P. Blaney
The Centre for Continuing Education
The University of British Columbia

The Application of Formative and Summative Evaluation Techniques to Short Term Adult Education Courses*

Summary

This paper identifies a problem concerning the practice of adult education, and describes one person's response to that problem. The problem is the lack of systematic evaluation of instruction in our field and, more particularly, the absence of sufficient evaluation literature that is ~~most~~ both useful and practical to those who earn their living in continuing education. The response is a developmental project aimed at producing for practitioners a set of materials that can be used in developing and evaluating instruction. The nature of these materials and how some practitioners have been trained to use them will be described.

The purpose of this paper is to generate discussion about whether the identified problem is real and accurately described, and how my particular response to it, if adequate, might be improved.

* Parts of this paper have been adapted from "A Guide to Developing and Evaluating Instruction in Continuing Education: Basic Concepts and Procedures", copyright by J.P. Blaney, 1972.

Problem

Many persons choose to work in continuing education because they are aware that the programs they organize have the potential of bringing about important changes within individuals and within our social institutions. Those who are organizing courses for practicing professionals, for adults curious and concerned about the society in which they live, or for a company's employees, often gain satisfaction in knowing that what their adult students may learn can be utilized almost immediately in various real-life situations. Generally speaking, however, continuing education courses do not reach their potential effectiveness. Or, when such courses are considered effective, judgments are based on faith, not evidence. For several reasons, existing technical knowledge concerning the improvement and measurement of instructional effects is infrequently employed in our field of work. One reason for this situation, I believe, is that the standard and most reputable references on the development and evaluation of instruction are not entirely appropriate for the unique nature of continuing education activities.

Continuing education differs in some important ways from public school and college education. Relatively speaking, continuing education is a marginal activity for institutions, with few resources for course development. Further, participation in education by adults is a part-time and generally a voluntary activity. Courses for adult students are often planned with these students, or at least with their perceived needs and

interests as the primary determiners of objectives and content. Also, there is no guaranteed student group for continuing education courses; if the prospective learners do not like some particular course, they simply do not attend, or drop out.

A significant consequence of the special nature of continuing education is that there is very little in the way of large scale curriculum development. Perhaps there should be, and will be some day, but presently there is not. There is nothing in continuing education to compare with the public school and college introductory courses in mathematics, national history, chemistry, and so on. Projects of the size and scope of the PSSC physics program* or Sesame Street are simply not found in North American adult education. While there are some attempts at developing packaged instructional materials for adults, these represent an extremely small fraction of current continuing education activities. Typically, continuing education courses are relatively short, one-shot events whose objectives are locally determined.

With its focus the curricula of public schools and colleges, the prominent course development and evaluation literature concerns major curriculum projects. The theories and methodologies of this literature are aimed not only at improving courses through protracted tryout-revision cycles, but also at describing the effects due only to a certain course and comparing one course against others. Evaluation-research designs utilizing control group situations are explicitly or implicitly suggested.

---

* A four-part introductory physics course for high schools prepared by the Physical Science Study Committee, Education Services Inc., Newton, Massachusetts.

Recommended procedures are complex and costly. Most persons in continuing education who have consulted this literature find its methodologies of great interest, but largely too elaborate and impractical. But, because it is our professional responsibility to maximize and demonstrate our effectiveness to both adult students and to those who support our institutions, we cannot ignore what this literature has to offer.* Rather, we should extract from it those tools we most need and can afford, and adapt them to the situations in which we are involved.

It is suggested here that the primary need respecting evaluation in continuing education is for reasonably uncomplicated procedures which can be used to improve any single course, and to describe what students know or feel at the conclusion of that course. Though still not simple or inexpensive, single-group evaluation procedures that focus on program development and less on certification are far more realistic for continuing education than the more elaborate procedures developed for large scale projects and packaged curriculum materials. Also, as noted by Guba (1969), evaluations incorporating experimental designs prevent rather than promote changes in instruction since treatments must remain untampered. Furthermore, the results of such evaluations have tended to be too insignificant to be of any practical use.

---

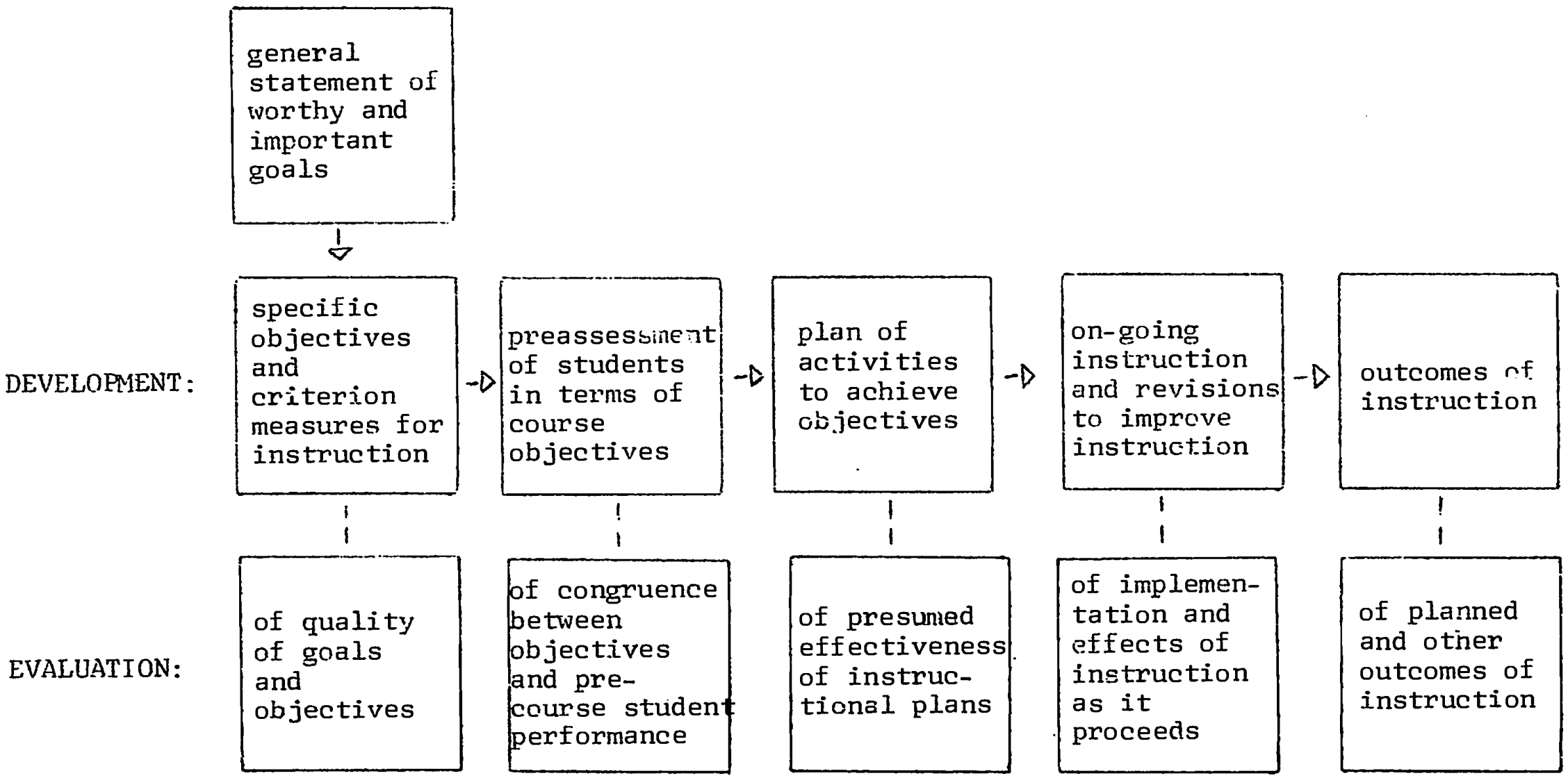
* Selected references appended.

A Model and Operating Procedures for the Systematic Development and Evaluation of Instructional Programs: One Response to the Problem

Following from the above discussion, we need effective but practical models with operating procedures for developing and evaluating instruction in continuing education. Further, these models should be constructed so that users if they wish may employ only selected elements and their related procedures. One model of this kind is proposed and displayed on page six. Its emphasis is the improvement of program quality and effectiveness, rather than program validation. The model's focus is the learner --what he is expected or hopes to do, or feel, what he can do before instruction, what activities are likely to assist him in achieving his goals, how he performs during instruction, and what he can do after instruction.

Most models that represent attempts to simplify the course development and evaluation process usually have the following sequence of components: specific objectives; preassessment of learners; instructional activities; and evaluation. The unfortunate implication of such models is that evaluation is an end-of-instruction task, or that the major purpose of evaluation is program certification. As suggested on page six, however, evaluation activities of various kinds should occur throughout the whole process of developing instructional programs. In other words, evaluative operations are required to make intelligent decisions about the appropriateness and worth of goals, the design and management of instructional procedures, and the impact

A Model for the Systematic Development and Evaluation of Instructional Programs*



* Implicit in this model is recycling at any point.

of instruction. Viewed this way, evaluation becomes an integral part of planning and improving every phase of instruction.

Drawing upon major works concerning formative and summative evaluation, criterion-referenced instruction, multiple criterion measures, etc. (see References), several technical papers describing basic evaluation concepts and techniques were formulated. Subsequently, with these papers as a basis, 27 explicit operational procedures were formulated in order to provide clear guidelines for the implementation of the above model. The technical papers and operating procedures, having undergone several revisions, were then combined into a 100 page guidebook-monograph for practitioners in continuing education. The monograph will be revised further on the basis of several field tests.

The 27 operational procedures for using the evaluation model cannot be detailed here, but the following summary indicates how they are categorized and the evaluation concepts and techniques to which they relate.

1. Selection and Justification of Instructional Goals: collection of an array of needs/goals; setting priorities.

2. Formulation of Instructional Objectives and Criterion Measures: specification of cognitive and affective objectives; instructional effectiveness standards; entry behavior; sequencing objectives; criterion measures (multiple and unobtrusive); unmeasurable objectives; item sampling; using preassessment data.

3. Development of an Instructional Plan: tryouts--whole or partial; principles of instructional psychology and adult learning; unit construction of instructional plan; diagnoses of instructional effectiveness and individual learner progress; remedial instruction; selecting media; instructional facilities.

4. Management of Instruction: monitoring the implementation of the instructional plan; revisions to on-going instruction.

5. Terminal Evaluation and Follow-up: immediate and delayed collection of outcome data; unanticipated outcomes; use of outcome data; costs of results; evaluating the course development process and operating procedures; evaluation report.

Training Practitioners to Use the Model and Operating Procedures; Evaluating the Training Materials

Using the materials noted above, three 3-day courses have been conducted in the past year. In general performance terms, the objectives for the most recent course were as follows:

(terminal, overall course objective)

- Upon completing the course participants will be able to use a set of given guidelines (27 operational procedures) in developing and evaluating instruction.

(enroute course objectives)

- Describe the three major elements (questions for which information should be obtained) of an overall evaluation of an instructional program.

- Reconstruct a particular model for the systematic development and evaluation of an instructional program; describe the relationships among the components of this model.
- Explain the use of the Q-sort technique as a tool for establishing priorities among competing instructional goals.
- Construct explicit objectives and criterion tests for evaluating the effectiveness of instruction.
- Describe the following terms; given situations which call for the use of any of these terms or operations, select the one(s) applicable:
    a) formative evaluation
    b) summative evaluation
    c) criterion-referenced measurement
    d) norm-referenced measurement
    e) content validity
    f) item sampling.
- Provide examples of test items or testing situations that fall into student product and student behavior classifications under both natural and manipulated conditions.
- Describe three important instructional variables and how they can be incorporated into a plan of instruction; construct a simple plan of instruction which incorporates these variables.

Objectives for the first two courses were similar to those given above, and differed mainly with respect to the nature of the terminal objective since the 27 operating procedures were

available only for the third course. Participants were not screened for any of the courses, however none of the students had previous training in evaluation. Each course included the following instructional techniques: reading assignments; discussions of readings; practice exercises related to the objectives; and demonstrations of recommended procedures. In summary form, the varying characteristics and results of the three courses were as follows:

	participants	materials used	testing procedures	overall results
course I	16 practitioners	1st version of technical papers	pretest and posttest	62% of students attained 75% or more objectives; 37% of students attained 80% or more
course II	28 graduate students	revised technical papers	pretest, posttest, one set of formative tests	82% of students attained 80% or more objectives
course III	35 practitioners	guidebook integrating papers plus 27 specific procedures	pretest, posttest and 2 sets of formative tests	90% of students attained 85% or more objectives

For additional information, a report on the first-conducted course is appended.

Further developmental work will be undertaken. A follow-up study of the participants in the third course will be made to establish whether the evaluation guidelines are in fact used on the job and, if not, why not. Deficiencies in the current

training materials also will be identified. In addition, a revised set of training materials (the 100 page monograph plus practice exercises) will be tested to roughly determine the extent to which the materials, by themselves, can assist 20 or more practitioners to attain the kind of objectives described earlier.

SELECTED REFERENCES

American Institutes for Research. Evaluative Research - Strategies and Methods. Pittsburgh, Pennsylvania: American Institutes for Research, 1970.

Ausubel, David P. The Psychology of Meaningful Verbal Learning. New York: Grune and Stratton, Inc., 1963.

Ausubel, David P. Learning Theory and Classroom Practice. The Ontario Institute for Studies in Education. Bulletin No. 1, 1967.

Baker, Eva L. Defining Content for Objectives. (A filmstrip-tape program) Los Angeles: Vimcet Associates, 1968.

Blaney, John P. and McKie, Douglas. "Knowledge of Conference Objectives and Effect upon Learning." Adult Education, XIX (1969), 98-105.

Blaney, John P. "The Effects of Differences in Training on Teacher Construction of Instructional Plans." Doctoral Dissertation, UCLA, 1970.

Blaney, John P. and Overing, Robert L. The Improvement of Instruction. Interim Position Paper for the Human Resources Research Council, Edmonton, Alberta, 1970.

Block, James H., ed. Mastery Learning Theory and Practice. New York: Holt, Rinehart and Winston, Inc., 1971.

Bloom, Benjamin S.; Hastings, J. Thomas; Madaus, George F. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill Book Company, 1971.

Briggs, L.J. Sequencing of Instruction in Relation to Heirarchies of Competence. Pittsburgh: American Institutes for Research, 1968. Monograph No. 3.

Dalis, Gus. "The Effect of Precise Objectives upon Student Achievement in Health Education." Doctoral Dissertation, University of California, 1969.

Downey, Lawrence Wm. The Task of Public Education - The Perceptions of People. Chicago: Midwest Administration Center, The University of Chicago, 1964.

Feldhusen, John F. and Treffinger, Donald J. "Psychological Background and Rationale for Instructional Design." Educational Technology, October, 1971, 21-24.

Gibb, Jack R. "Learning Theory in Adult Education." Handbook of Adult Education in the United States. Chicago: Adult Education Association of the U.S.A., 1960.

Glaser, Robert. Evaluation of Instruction and Changing Educational Models. Occasional Report No. 13, Center for the Study of Evaluation of Instructional Programs. Los Angeles: University of California, 1968.

Goodlad, John I. The Development of a Conceptual System for Dealing with Problems of Curriculum and Instruction. Los Angeles: Institute for the Development of Educational Activities, 1966.

***

Harty, L.D. and Monroe, B.P. Objectives for Instructional Programs. Long Beach, California: Insgroup, Inc., 1970.

Hively, Wells, II. "Introduction to Domain-Referenced Achievement Testing." Symposium Presentation, American Educational Research Association, Minneapolis, Minnesota, March 2-6, 1970.

*

Husek, T.R. and Sirotnik, K. Item Sampling in Educational Research. Occasional Report No. 2, Center for the Study of Evaluation of Instructional Programs. Los Angeles: University of California, 1967.

Kerlinger, Fred N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc., 1964.

Lumsdaine, A.A. "Assessing the Effects of Instructional Programs." Teaching Machines and Programmed Learning, II: Data and Directions. Edited by R. Glaser. Washington, D.C.: National Educational Association, 1965.

Mager, Robert F. Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers, 1962.

Mager, Robert F. and Beach, Kenneth M., Jr. Developing Vocational Instruction. Palo Alto, California: Fearon Publishers, 1967.

Mager, Robert F. Developing Attitude Toward Learning. Palo Alto, California: Fearon Publishers, 1968.

**

Maguire, Thomas O. "Value Components of Teachers' Judgments of Educational Objectives." AV - Communication Review, XVI (Spring, 1968), 63-86.

Markle, S.M. "Empirical Testing of Programs." Program Instruction. Edited by P.C. Lange. Sixty-Sixth Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1967.

Millman, Jason. "Reporting Student Progress: A Case for a Criterion-Referenced Marking System." Phi Delta Kappan, (December, 1970), 226-230.

Popham, W. James and Baker, Eva L. Rules for the Development of Instructional Products. Inglewood, California: Southwest Regional Laboratory for Educational Development, 1967.

Popham, W. James and Baker, Eva L. Systematic Instruction. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970a.

Popham, W. James and Baker, Eva L. Establishing Instructional Goals. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970b.

Popham, W. James. An Evaluation Guidebook. Los Angeles: The Instructional Objectives Exchange, 1971a.

Popham, W. James. Criterion-Referenced Measurement - An Introduction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971b.

Osgood, C.E.; Suci, G.J.; and Tannebaum, P.H. The Measurement of Meaning. Urbana, Illinois: University of Illinois Press, 1957.

Rogers, Everett M. Diffusion of Innovations. New York: The Free Press of Glencoe, 1962.

Scriven, Michael. "The Methodology of Evaluation." Perspectives on Curriculum Evaluation. Chicago: Rand McNally and Company, 1967.

Stufflebeam, Daniel I. et al. Educational Evaluation and Decision Making. Itasca, Illinois: F.E. Peacock Publishers, Inc., 1971.

Suchman, Edward A. Evaluative Research. New York: Russell Sage Foundation, 1967.

Stake, Robert E. "Objectives, Priorities, and Other Judgement Data." Review of Educational Research - RER, Vol. 40, No. 2 (April, 1970), 181-212.

Stake, Robert E. and Gooler, Dennis. "Measuring Educational Priorities." Educational Technology, Vol. XI, No. 9 (September, 1971), 44-48.

Verner, Coolie, and Booth, Alan. Adult Education. New York: The Center for Applied Research in Education, Inc., 1964.

Webb, Eugene,; Campbell, Donald T.; Schwartz, Richard D.; and Sechrest, Lee. Unobtrusive Measures - Nonreactive Research in the Social Sciences. Chicago: Rand McNally and Company, 1966.

Wesman, A.G. "Writing the Test Item." Educational Measurement. 2nd edition. Washington, D.C.: American Council on Education, 1971 81-129.

*Hovland, C.I.; Lumsdaine, A.A.; and Sheffield, F.D. Experiments on Mass Communication. Princeton, New Jersey: Princeton University Press, 1949.

**Mager, Robert F. and McCann, J. Learner Controlled Instruction. Palo Alto, California: Varian Associates, 1961.

***Guba, E.G. "The Failure of Education Evaluation." Educational Technology, 1969, 9 (5), 29-38.

Evaluation of Short-Term Extension Courses:

A Case Study

J.P. Blaney

and

Gary Dickinson

The University of British Columbia

1971

Few attempts are made to evaluate systematically the effectiveness of continuing education short courses and conferences. Such studies are generally regarded by program administrators as costly luxuries offering scant benefits because short-term courses are rarely repeated. In addition, the programs may be poorly designed, lacking clearly stated objectives so that there is no basis for a systematic evaluation. When evaluation of short-term programs is attempted, it usually consists of participant reactions to the program or intuitive judgments by the instructor or administrator. Those types of evaluation, however, are of little use in assessing the learning that resulted from the program or in improving the design of subsequent programs of a similar nature.

In order to encourage the increased use of evaluation procedures in continuing education short courses, a three day course on evaluation was offered in February, 1971 by the Centre for Continuing Education of The University of British Columbia in cooperation with the Canadian Association of Departments of Extension and Summer Schools. Sixteen program administrators from six university extension departments in Western Canada completed the course which was designed to illustrate the principles that were advocated. This article describes the evaluation process that was followed and the results that were obtained from using various measures of the effectiveness of the program.

DESIGN FOR EVALUATION

The design used for evaluating the short course was relatively simple, involving one group which was given a pretest and a posttest on the instructional objectives for the program. This design provides information about learner performance on instructional objectives both before and after instruction is given. The cost of using this procedure is fairly low, so the design can be used readily in other short courses at no great expense once the objectives are stated clearly.

The course on evaluation was planned, taught, and evaluated following the six steps noted briefly below. Those stages provide the structure for the remainder of this case study as each item is discussed in detail in the following sections.

1. Formulation of precise program objectives.
2. Construction of a test to measure student performance on the program objectives.
3. Pretest of learners to determine their entry behaviour with respect to the program objectives.
4. Provision of instruction relevant to the program objectives.
5. Posttest of learners to determine their terminal behaviour with respect to the program objectives.
6. Analysis of pretest and posttest data together with other indices of effectiveness to make an overall assessment of the program.

1. PROGRAM OBJECTIVES

Eight behavioural objectives were formulated and circulated to the intended participants several weeks prior to the

course. The objectives were modified slightly shortly before the course began and discussed with the participants on the first morning of the course. At the conclusion of the course, the participants were expected to be able to:

1. Given a hypothetical evaluation task, indicate the three basic questions for which information should be obtained.

2. Describe the essential components of a behavioural instructional model.

3. Given the basic procedures involved in formative and summative evaluation, discuss their application to a particular program with which they are concerned.

4. Given an evaluation question(s) concerning the effectiveness of instruction, choose or name the evaluation research design that will provide the required information at a minimum cost.

5. Describe the following terms; given situations which call for the use of any of these terms or operations, select the ones(s) applicable.

   - formative (developmental) evaluation
   - summative (terminal) evaluation
   - criterion-referenced measurement
   - norm-referenced measurement
   - content validity
   - item sampling

6. Generate a variety of observable behavioural indices (both cognitive and affective) for assessing the attainment of a non-operationally stated general objective.

7. Supply examples of behavioural indices that fall into student product and student behaviour classifications under both natural and manipulated conditions.

8. Given an evaluation report, identify its major strengths and weaknesses.

2. TEST CONSTRUCTION

Test items were constructed for each course objective except number 8. A total of ten test items were prepared, with single items for four objectives and two items for three objectives. The items required several responses so that the maximum score attainable was 56 points. The number of test points per objective ranged from 4 points (Objectives 2 and 7) to 18 points (Objective 5) with a mean of 8 points per objective Comments made by the participants suggested that three items might not have been clearly understood: two because of unfamilia terminology and one because of ambiguous wording. Only one form of the test was constructed so that the same instrument was used for both pretest and posttest.

3. ENTRY BEHAVIOUR

The untimed pretest was administered without prior warnin to the participants at the beginning of the course. Pretest scores ranged from 2 to 14 points with a mean of 8.56 and a standard deviation of 3.632. (Table 1) As the total scores on the pretest were quite low, no detailed analysis was made of entry behaviour in relation to the specific course objectives.

In addition to the pretest scores, other information was obtained about the possible entry behaviour of the participants through their completion of detailed preregistration forms. Those questionnaires indicated that none of the participants had received training in evaluation and none were using evaluation procedures in their programs except for post-course participant reactions or intuitive assessment.

4. INSTRUCTION

The course was planned quite flexibly in order to provide as much individual instruction and consultation as possible. Morning sessions were designed primarily for the acquisition of information and the afternoon sessions sought the application of knowledge. A set of instructional materials consisting of definitions, working papers, and articles was provided to each participant. Most of those materials were read silently, followed by question and discussion periods. A small library of materials on evaluation was made available so that the participants could pursue their own reading interests under guidance from the instructors. Some ad hoc groups were formed in the afternoons to discuss evaluation topics and materials related to specific program areas such as business administration and agricultural extension.

Some instructional activities were conducted that did not relate directly to the prespecified course objectives. Three such activities were a demonstration of the Q-sort technique, a discussion of a paper on the application of evaluation techniques to university continuing education courses, and viewing and discussion of a two-projector slide and tape presentation.

5. TERMINAL BEHAVIOUR

The posttest was untimed and administered at the conslusion of the course. Scores ranged from 29 to 49 points with a mean of 40.63 and a standard deviation of 5.289. A t-test indicated that the difference between pretest and posttest mean scores was

TABLE 1

PARTICIPANT PRETEST AND POSTTEST RESULTS

Participant No.	Pretest Score	Posttest Score
1	8	44
2	13	49
3	8	40
4	2	44
5	9	43
6	10	35
7	3	43
8	14	47
9	14	47
10	9	35
11	7	41
12	7	38
13	4	38
14	10	41
15	7	29
16	12	36
Mean	8.56	40.63

statistically significant beyond the .001 level. A rank correlation coefficient of .28 was obtained between the participants' pretest and posttest scores. That coefficient is not statistically significant at the .05 level, indicating that there was no association between scores on the two tests.

Learner performance on the posttest was analyzed for each behavioural objective that was measured. As Table 2 shows, the mean score as a percentage of the maximum score attainable varied from 44 to 93 percent. Objectives 1, 2, and 7 had the highest mean posttest percentages, and the instructors judged those objectives as having had the most provisions for practice during the course. Only objectives 4 and 3 which had mean scores of 44 and 55 per cent fell below the 75 per cent mark. The material related to objective 4 was discussed very briefly and no direct practice was provided. The scores on the test items related to objective 3 were influenced by the awarding of additional points for multiple responses when the questions did not clearly indicate that more than one response was required.

6. ADDITIONAL INDICES

The Kropp-Verner Attitude Scale for evaluating meetings was administered to the participants at the conclusion of the course. That scale has a score range of 1 (most favourable) to 11 (least favourable), with a score of 6 indicating an indifferent attitude. The attitude scores for the evaluation course ranged

TABLE 2

MEAN POSTTEST SCORE BY INSTRUCTIONAL OBJECTIVE

Objective No.	Maximum Score Attainable	Mean Score	Mean Score as Percentage of Maximum Score
1	6	5.56	93
2	4	3.56	89
3	8	4.38	55
4	10	4.38	44
5	18	14.63	81
6	6	4.56	76
7	4	3.56	89
Total	56	40.63	73

from 2.79 to 7.12 with a mean of 4.03 and a standard deviation of 1.023, which suggests that the course was regarded fairly favourably by the participants. Rank correlation coefficients were computed for attitude scale score in relation to pretest (.21) and posttest score (.18), but the coefficients obtained were not statistically significant. Attitudes toward the course therefore bore no relationship to the entry or terminal performance of the participants in relation to the course objectives.

Objective 8 was not examined by the test, however, two evaluation reports were analyzed by the group and this performance was judged to be satisfactory by the instructors in that the level of discussion would not have been likely to occur prior to the course.

The instructors both noted that the frequency and quality of oral participation by the learners improved throughout the course. The proportion of comments and questions dealing with the application of evaluation concepts increased while questions pertaining to information and clarification decreased.

Written comments about the course were sought from the participants. Most of the small number that were given indicated a fairly high degree of satisfaction. One of the original seventeen participants did not attend the second half of the course.

In conclusion, the instruction provided to the participants in the short course on evaluation led to satisfactory learner performance with respect to six of the eight objectives that were

formulated prior to the course. Considerable gains in learner performance were noted and the participants were generally satisfied with the course.